

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 16:23:18 ; Search time 176 Seconds
(without alignments)
320.049 Million cell updates/sec

Title: US-10-823-203-3
Perfect score: 545
Sequence: 1 MSLKSDVFAKIAKRLKLESID.....EVDGQVELIFLLEPFIASLK 110

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	545	100.0	110	2 Q86PR3	Q86pr3 aedes aegypti
2	540	99.1	110	2 Q86PMO	Q86pmo culex quinquefasciatus
3	536	98.3	110	2 Q816X4	Q816x4 aedes aegypti
4	534	98.0	110	2 Q81T68	Q81t68 aedes aegypti
5	467	85.7	110	2 Q81T65	Q81t65 culex quinquefasciatus
6	459	84.2	110	2 Q816X3	Q816x3 anopheles gambiae
7	456.5	83.8	111	2 Q81T63	Q81t63 culex quinquefasciatus
8	452.5	83.0	111	2 Q81T64	Q81t64 culex quinquefasciatus
9	452.5	83.0	111	2 Q81T66	Q81t66 culex quinquefasciatus
10	452.5	83.0	111	2 Q81T67	Q81t67 aedes aegypti
11	452.5	83.0	111	2 Q7KAU2	Q7kau2 aedes aegypti
12	439	80.6	111	2 Q7Q8A4	Q7pq84 anopheles gambiae
13	172	31.6	164	2 Q7PHX9	Q7phx9 anopheles gambiae
14	172	31.6	728	2 Q7Q8E6	Q7q8e6 anopheles gambiae
15	161.5	29.6	115	2 Q6X112	Q6x112 drosophila
16	161.5	29.6	115	2 Q9VY92	Q9vy92 drosophila
17	144	26.4	107	2 Q9VE19	Q9ve19 drosophila
18	130.5	23.9	121	1 NLTP_BOVIN	NLTPbovin bos taurus
19	129.5	23.8	547	1 NLTP_CHICK	NLTPchick gallus gallus
20	129.5	23.8	547	1 NLTP_RAT	NLTPrat rattus norvegicus
21	128.5	23.6	547	1 NLTP_MOUSE	NLTPmouse mus musculus
22	127.5	23.4	547	1 NLTP_HUMAN	NLTPhuman homo sapiens
23	125.5	23.0	547	1 NLTP_RABIT	NLTPrabbit oryctolagus
24	123.5	22.7	737	2 Q28956	Q28956 sus scrofa
25	122.5	22.5	106	2 Q19066	Q19066 sus scrofa
26	120.5	22.1	735	2 Q6RF26	Q6rf26 equus caballus
27	118.5	21.7	736	1 DHB4_HUMAN	DHB4human h. peroxisom
28	116.5	21.4	536	1 Q640H2	Q640h2 xenopus laevis
29	115.5	21.2	169	2 Q7PQP5	Q7pqp5 anopheles gambiae
30	115.5	21.2	169	2 Q7PTV8	Q7ptv8 anopheles gambiae
31	114.5	21.0	735	2 Q42484	Q42484 gallus gallus

32 113.5 20.8 734 2 P70523 P70523 rattus norvegicus
33 113.5 20.8 735 1 DHB4_RAT DHB4_rattus norvegicus
34 113.5 20.8 735 2 O70529 O70529 cavia porcellus
35 113.5 20.8 735 2 P70540 P70540 rattus norvegicus
36 112.5 20.6 725 2 Q98TA2 Q98ta2 brachydanio rerio
37 112.5 20.6 725 2 Q8AYH1 Q8ayh1 brachydanio rerio
38 110.5 20.3 725 2 Q6NZW5 Q6nzw5 brachydanio rerio
39 110 20.2 741 2 Q6GMC3 Q6gmc3 xenopus laevis
40 108.5 19.9 538 2 Q6P4V5 Q6p4v5 brachydanio rerio
41 108.5 19.9 736 2 Q68V19 Q68v19 bos taurus
42 107.5 19.7 735 1 DHB4_MOUSE DHB4_mus musculus
43 107.5 19.7 735 2 Q9DBM3 Q9dbm3 m. musculus
44 104.5 19.2 118 1 NLTI_CAERL NLTI_caenorhabditis elegans
45 100 18.3 436 2 Q21481 Q21481 caenorhabditis elegans

ALIGNMENTS

RESULT 1
Q86PR3 PRELIMINARY; PRT: 110 AA.
ID Q86PR3
AC Q86PR3
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Sterol carrier protein 2.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rockefeller;
RX MEDLINE=22431721; PubMed=12542635;
RA Krebs K.C., Lan Q.;
RT "Isolation and expression of a sterol carrier protein-2 gene from the yellow fever mosquito, Aedes aegypti.";
RL Insect Mol. Biol. 12:51-60(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rockefeller;
RA Krebs K.C., Brzosa K.L., Lan Q.;
RT "Use of Subtracted Libraries and Macroarray to Isolate Developmentally Specific Genes from the Mosquito, Aedes aegypti.";
RL Insect Biochem. Mol. Biol. 0:0-0(2003).
DR EMBL; AY190283; AA034708.1; --
DR PDB; 1PZ4; X-ray; A1-110.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
SQ SEQUENCE 110 AA; 12272 MW; 5A79FC1CC20298CE CRC64;
Query Match 100.0%; Score 545; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLKSDVFAKIAKRLKLESIDPANRQVEHYKFRITQGGKVVKNWMDLKNVXLVESDDAA 60
Db 1 MSLKSDVFAKIAKRLKLESIDPANRQVEHYKFRITQGGKVVKNWMDLKNVXLVESDDAA 60
QY 61 EATLTMEDDIMPAGTGPALPAKMAQDKMEVDGQVELIFLLEPFIASLK 110
Db 61 EATLTMEDDIMPAGTGPALPAKMAQDKMEVDGQVELIFLLEPFIASLK 110
RESULT 2
Q86PMO PRELIMINARY; PRT: 110 AA.
ID Q86PMO
AC Q86PMO
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

DE Sterol carrier protein 2 variant 2.
OS Culex quinquefasciatus (Southern house mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culex.
OX NCBI_TaxID=7176;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Borovsky D., Morris T.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AX196002; AA043438.1; -.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
SQ SEQUENCE 110 AA; 12271 MW; 525976346088B84C CRC64;

Query Match 99.1%; Score 540; DB 2; Length 110;
Best Local Similarity 99.1%; Pred. No. 1.1e-42;
Matches 109; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKSDVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
DB 1 MSLKSNVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
QY 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110
DB 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110

RESULT 3
Q8I6X4
ID Q8I6X4 PRELIMINARY; PRT; 110 AA.
AC Q8I6X4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sterol carrier protein 2 variant 2.
OS Aedes aegypti (Yellow fever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole mosquito;
RA Borovsky D., Thermote L.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF539995; AANI6392.1; -.
DR HSSP; P51659; 1IKT.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
SQ SEQUENCE 110 AA; 12272 MW; 6939A03CCAA01AC6 CRC64;

Query Match 98.3%; Score 536; DB 2; Length 110;
Best Local Similarity 98.2%; Pred. No. 2.7e-42;
Matches 108; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKSDVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
DB 1 MSLKSNVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
QY 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110
DB 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110

RESULT 4
Q8IT68
ID Q8IT68 PRELIMINARY; PRT; 110 AA.
AC Q8IT68;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Sterol carrier protein 2 variant 2.
OS Aedes aegypti (Yellow fever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Borovsky D., Thermote L.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF539988; AANI6385.1; -.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
SQ SEQUENCE 110 AA; 12303 MW; 494297717588B84D CRC64;

Query Match 98.0%; Score 534; DB 2; Length 110;
Best Local Similarity 97.3%; Pred. No. 4.1e-42;
Matches 107; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKSDVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
DB 1 MSLKSNVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
QY 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110
DB 61 EATLTWEDDIMPAGTGAMPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110

RESULT 5
Q8IT65
ID Q8IT65 PRELIMINARY; PRT; 110 AA.
AC Q8IT65;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sterol carrier protein 2 variant 2.
OS Culex quinquefasciatus (Southern house mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culex.
OX NCBI_TaxID=7176;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gut;
RA Borovsky D., Thermote L.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF539991; AANI6388.1; -.
DR HSSP; P51659; 1IKT.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
SQ SEQUENCE 110 AA; 12362 MW; 4A4EC73A9CC1D5CA CRC64;

Query Match 85.7%; Score 467; DB 2; Length 110;
Best Local Similarity 86.4%; Pred. No. 7.2e-36;
Matches 95; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSLKSDVEPAKTAKRLESIDPANRQVEHYVKFRITGGKVKVKNWMDLKNVKLVESDDAA 60
DB 1 MSLKSNVEPAKTAKRLENIDPANRQVQYVKFRITKDGKVKVKNWMDLKNVKLVESDGA 60
QY 61 EATLTWEDDIMPAGTGALPAKEMAAQDKNEVDGQVELIFLLEPFIASLK 110
DB 61 EATLTWEDDIMPAGTGAMPAKEMAAQDKNEVRGQVELIFLLEPFIASLK 110

RESULT 6
Q8I6X3
ID Q8I6X3 PRELIMINARY; PRT; 110 AA.
AC Q8I6X3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

[illegible]

DE	Sterol carrier protein 2.	GN	Name=ENSANGG0000018004;
OS	Aedes aegypti (YellowFever mosquito).	OS	Anopheles gambiae str. PEST.
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.	OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX	NCBI_TaxID=7159;	OX	NCBI_TaxID=180454;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	TISSUE=Gut;	RC	STRAIN=PEST;
RA	Borovsky D., Thernote L.;	RA	Anopheles Genome Sequencing Consortium;
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	RL	Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF539989; AANI6386.1; -.	CC	-!- CAUTION: The sequence shown here is derived from an
DR	GO: 0005498; F:sterol carrier activity; IEA.	CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
DR	InterPro; IPR003033; SCP2.	CC	preliminary data.
DR	Pfam; PF02036; SCP2; 1.	DR	EMBL; AAB0100879; EAA08376.2; -.
SQ	SEQUENCE 111 AA; 12491 MW; 9E28F2E1A0564E49 CRC64;	DR	GO: 0005498; F:sterol carrier activity; IEA.
		DR	InterPro; IPR003033; SCP2.
		DR	Pfam; PF02036; SCP2; 1.
		FT	NON TER 1
		SQ	SEQUENCE 164 AA; 17529 MW; 9B1DED25675ESA42 CRC64;
		Query Match	31.6%; Score 172; DB 2; Length 164;
		Best Local Similarity	33.6%; Pred. No. 3.5e-08;
		Matches	37; Conservative 23; Mismatches 46; Indels 4; Gaps 1;
QY	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLESDDAA 60	QY	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLESDDAA 60
Db	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLESDDAA 60	Db	1 MALKSDPVFERIAKRLSIDPNNRQVQVYKFRITGGKVVKNVMDLKNVLESDDAA 60
QY	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 110	QY	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 110
Db	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 111	Db	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 110
		RESULT 11	
Q7KAU2	PRELIMINARY; PRT; 111 AA.	Q7PHX9	PRELIMINARY; PRT; 164 AA.
ID	Q7KAU2	ID	Q7PHX9
AC	Q7KAU2; 0810E3;	AC	Q7PHX9;
DT	05-JUL-2004 (TREMBlrel. 27, Created)	DT	01-MAR-2004 (TREMBlrel. 26, Created)
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)	DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE	Sterol carrier protein 2 variant 1.	DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)
OS	Aedes aegypti (YellowFever mosquito).	DE	ENSANGP0000023849 (Fragment).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	GN	Name=ENSANGG0000011810;
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.	OS	Anopheles gambiae str. PEST.
OX	NCBI_TaxID=7159;	OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
RN	[1]	OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
RP	SEQUENCE FROM N.A.	OX	NCBI_TaxID=180454;
RC	TISSUE=Brain, and Whole mosquito;	RN	[1]
RA	Borovsky D., Thernote L.;	RP	SEQUENCE FROM N.A.
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	RC	STRAIN=PEST;
DR	EMBL; AF539994; AANI6391.1; -.	RA	Anopheles Genome Sequencing Consortium;
DR	EMBL; AF539987; AANI6384.1; -.	RL	Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR	InterPro; IPR003033; SCP2.	CC	-!- CAUTION: The sequence shown here is derived from an
DR	Pfam; PF02036; SCP2; 1.	CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
SQ	SEQUENCE 111 AA; 12519 MW; 938523F7D139700E CRC64;	CC	preliminary data.
		DR	EMBL; AAB01008944; EAA44362.1; -.
		DR	HSSP; P51659; I1K7.
		DR	GO: 0005498; F:sterol carrier activity; IEA.
		DR	GO: 0005498; F:sterol carrier activity; IEA.
		DR	InterPro; IPR003033; SCP2.
		DR	Pfam; PF02036; SCP2; 1.
		FT	NON TER 1
		SQ	SEQUENCE 164 AA; 17529 MW; 9B1DED25675ESA42 CRC64;
		Query Match	31.6%; Score 172; DB 2; Length 164;
		Best Local Similarity	33.6%; Pred. No. 3.5e-08;
		Matches	37; Conservative 23; Mismatches 46; Indels 4; Gaps 1;
QY	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLESDDAA 60	QY	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLES 56
Db	1 MSLKSDVEFAKIAKRLSIDPANROVHYHYKFRITGGKVVKNVMDLKNVLESDDAA 60	Db	49 VDLQSDAVFAGIKDRVAENEAKAINAVFLYKITSGGKVAKEWLDLKNVYVEGPVG 108
QY	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 110	QY	57 DDAEATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFI 106
Db	61 EATLTMEDDIMPFAITGALPAKEMAAQDKMEVDGQVE-LIFLLEPFIASLK 111	Db	109 GKGADTTMTIADGDMIELALGKLQPTAPMKGLKIAGNIMLAQKLAPLL 158
		RESULT 12	
Q7POS4	PRELIMINARY; PRT; 111 AA.	Q7POS4	PRELIMINARY; PRT; 111 AA.
ID	Q7POS4	ID	Q7POS4
AC	Q7POS4;	AC	Q7POS4;
DT	01-MAR-2004 (TREMBlrel. 26, Created)	DT	01-MAR-2004 (TREMBlrel. 26, Created)
DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)	DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE	ENSANGP0000020493.	DE	ENSANGP0000020493.

```

RESULT 14
Q7Q8B6
ID Q7Q8B6 PRELIMINARY; PRT; 728 AA.
AC Q7Q8B6
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE AGCP15453 (Fragment).
GN Name=agCG51704; ORFNames=ENSGANG00000011810;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAINS=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008944; EAA10131.1; -.
DR HSSP; P97852; IG26.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR InterPro; IPR002539; MacC_dehydratas.
DR InterPro; IPR00169; Pept_cys_acsite.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF01575; MacC_dehydratas; 1.
DR Pfam; PF02036; SCP2; 1.
DR PRINTS; PRO0081; GDRHDH.
DR PRINTS; PRO0080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
DR PROSITE; PS00639; THIO_L_PROTEASE_HIS; UNKNOWN_1.
KW Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 728 AA; 78295 MW; 252404154801D872 CRC64;

Query Match 31.6%; Score 172; DB 2; Length 728;
Best Local Similarity 33.6%; Pred. No. 1.6e-07;
Matches 37; Conservative 23; Mismatches 46; Indels 4; Gaps 1;

QY 1 MSLKSDVEVFAKIARLESIDPANRQVEHVYKFRITGGKVKVKNWMDLKNVKLVESDD--VES 56
Db 613 VDLQSDAVFQKIIDGLKDNKAKAVNGVFLYKITKDGKVAKEWLDLKNKAVYEGPVQG 672

QY 57 DDAEATLTMEDDIMPFAITGTGALPAKEAMAQDKMEVDGQVELIFLLEPFI 106
Db 673 GKGADTTWTIADGDMIELALGKLQPTAFPMKGKLIAGNIMLTQKLAPLL 722

RESULT 15
Q8XII2
ID Q8XII2 PRELIMINARY; PRT; 115 AA.
AC Q8XII2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Similar to Drosophila melanogaster CG11151 (fragment).
OS Drosophila yakuba (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=22887302; PubMed=14525923; DOI=10.1101/gr.1311003;
RA Domazet-Lozo T., Tautz D.;
RT "An evolutionary analysis of orphan genes in Drosophila.";
RL Genome Res. 13:2213-2219(2003).
DR EMBL; AY232020; AAR10043.1; -.
DR GO; GO:0005498; F:sterol carrier activity; IEA.
DR InterPro; IPR003033; SCP2.
DR Pfam; PF02036; SCP2; 1.
FT NON TER 1
FT NON TER 115
SQ SEQUENCE 115 AA; 12439 MW; BF1E490B24AD681A CRC64;

Query Match 29.6%; Score 161.5; DB 2; Length 115;
Best Local Similarity 34.9%; Pred. No. 2.3e-07;
Matches 38; Conservative 19; Mismatches 49; Indels 3; Gaps 1;

QY 1 MSLKSDVEVFAKIARLESIDPANRQVEHVYKFRITGGKVKVKNWMDLKNVKLVESDD-- 58
Db 1 MSLQSDAVFQKIIDGLKDNKAKAVNGVFLYKITKDGKVAKEWLDLKNKAVYEGPVQG 60

QY 59 -AAEATLTMEDDIMPFAITGTGALPAKEAMAQDKMEVDGQVELIFLLEPFI 106
Db 61 IKVDTTLTVADEDMVDIALGKLNQQAAPFMKGKLIAGNIMLTQKLAPLL 109

Search completed: February 28, 2005, 16:40:04
Job time : 178 secs

```

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 16:22:28 ; Search time 163 Seconds
(without alignments)
261.004 Million cell updates/sec

Title: US-10-823-203-3

Perfect score: 545
Sequence: 1 MSLKSDVEFAXIARLESID.....EVDGQVELFLLEPFIASLK 110

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161.5	29.6	115	4	ABB65491 Drosophil
2	144	26.4	107	4	ABB61449 Drosophil
3	129.5	23.8	547	7	ADD47206 Rat Prote
4	128.5	23.6	547	5	ABB57301 Mouse lac
5	127.5	23.4	143	7	ADJ70149 Human hea
6	127.5	23.4	547	7	ADJ71194 Human hea
7	127.5	23.4	547	8	ABM80093 Tumour-as
8	118.5	21.7	735	2	AAW16329 Human hos
9	118.5	21.7	735	4	AAW16329 Human hos
10	118.5	21.7	736	4	AAW16329 Human mul
11	118.5	21.7	736	4	AAW16329 Human mul
12	118.5	21.7	736	5	ABG96550 Human sno
13	118.5	21.7	736	7	ADG61951 Human pro
14	118.5	21.7	736	7	ADG61947 Human pro
15	118.5	21.7	736	7	ADG60838 Human pro
16	118.5	21.7	752	8	ADG66395 Human pro
17	118.5	21.7	752	8	ADG66737 Human pro
18	113.5	20.8	735	7	ADG60836 Rat Prote
19	113.5	20.8	735	7	ADG61945 Rat Prote
20	113.5	20.8	735	7	ADG61949 Rat Prote
21	104	19.1	740	4	AAU32847 Novel hum
22	100	18.3	436	8	ADN24158 Bacterial
23	96	17.6	172	5	ABG60202 Human DIT
24	95.5	17.5	544	4	ABB65056 Drosophil
25	92.5	17.0	211	6	ABO00589 Novel hum

26	90	16.5	412	4	ABB61661	Abb61661 Drosophil
27	83	15.2	203	5	ADK36946	Adk36946 Novel hum
28	83	15.2	203	6	ABO00851	ABO00851 Polypepti
29	83	15.2	278	4	AAU23020	AAU23020 Novel hum
30	83	15.2	278	4	ABB10251	Abb10251 Human CDN
31	83	15.2	278	4	AAU18466	AAU18466 Human end
32	83	15.2	278	5	ABP66838	ABP66838 Human pol
33	83	15.2	345	7	ADJ70022	ADJ70022 Human hea
34	83	15.2	357	4	AAU18345	AAU18345 Human end
35	83	15.2	406	4	ABP37971	ABP37971 Human GS9
36	83	15.2	418	4	ABG84367	ABG84367 Amino aci
37	83	15.2	418	4	AAH81260	AAH81260 Human AFP
38	83	15.2	418	5	AAU76223	AAU76223 Human 216
39	83	15.2	418	5	ABU65161	ABU65161 Human NOV
40	83	15.2	418	8	ADN61973	ADN61973 Human can
41	83	15.2	418	8	ADQ15098	ADQ15098 Human can
42	83	15.2	422	4	ABP37972	ABP37972 Human GS9
43	81	14.9	71	5	ABP01408	ABP01408 Human ORF
44	81	14.9	154	5	ABP58804	ABP58804 Membrane-
45	80	14.7	926	8	ADR31493	ADR31493 Phosphoen

ALIGNMENTS

RESULT 1
ABB65491
ID ABB65491 standard; protein; 115 AA.
XX
AC ABB65491;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 23265.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
(PEKE) PE CORP NY.
XX
Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-658860/75.
XX
N-PSDB; ABL09594.
XX
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signaling and cell-cell
interactions.
XX
Dislosure; SEQ ID NO 23265; 21pp + Sequence Listing; English.
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
ABB72072). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences

```
SQ Sequence 115 AA;
Query Match          29.6%; Score 161.5; DB 4; Length 115;
Best Local Similarity 34.9%; Pred. No. 2.9e-10;
Matches 38; Conservative 19; Mismatches 49; Indels 3; Gaps 1;

QY 1 MSLSKDEVFQKIADGKLENEAKAKAVNGVFLYKITDKGKAKVETLDCNKAAYEGPAQG 60
DQ 1 MSLSQDAVFQKIADGKLENEAKAKAVNGVFLYKITDKGKAKVETLDCNKAAYEGPAQG 60

QY 59 -AAEATLTWEDDIFGAITGALPAKEMAQDKMEVDGQVELIFLLEPFI 106
DQ 61 IKVDTTLTVAEDMDVIALGKLNPOAFAFMKGLKIAGNIMLTQKLAPLL 109

RESULT 2
ID ABB61449
XX ABB61449 standard; protein; 107 AA.
AC ABB61449;
XX ABB61449;
DT 26-MAR-2002 (first entry)
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 11139.
XX Drosophila melanogaster polypeptide; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX WO200171042-A2.
PN 27-SEP-2001.
XX 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2001; 2001WO-US009231.
PF 23-MAR-2000; 2000US-0191637P.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX 11-JUL-2000; 2000US-00614150.
PA (PEKE ) PE CORP NY.
XX (PEKE ) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656960/75.
DR N-PSDB; ABL05552.
XX N-PSDB; ABL05552.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
PS Disclosure; SEQ ID NO 11139; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 107 AA;

Query Match          26.4%; Score 144; DB 4; Length 107;
Best Local Similarity 31.4%; Pred. No. 2.6e-08;
Matches 33; Conservative 25; Mismatches 41; Indels 6; Gaps 3;

QY 3 LKSDEVFQKIADGKLENEAKAKAVNGVFLYKITDKGKAKVETLDCNKAAYEGPAQG 61
DQ 1 MKSDIIEKIRNKLESDPARVTVNTFQNFDTADGNLIKSMALDIYE----GSATSYD 56
```

```
QY 62 ATLTMEDDIFGAITGALPAKEMAQDKMEVDGQVELIFLLEPFI 105
DQ 57 AQTTSIDSEDFYLVGTQKTFQEVLOQEKAKIDGDEAINKMLEKF 101

RESULT 3
ID ADD47206
XX ADD47206 standard; protein; 547 AA.
AC ADD47206;
XX ADD47206;
DT 02-DEC-2004 (revised)
DQ 29-JAN-2004 (first entry)
XX 29-JAN-2004 (first entry)
DE Rat Protein AAA41726, SEQ ID NO 12900.
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.
XX Unidentified.
XX WO2003016475-A2.
PN 27-FEB-2003.
XX 27-FEB-2003.
PD 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2002; 2002WO-US025765.
PF 14-AUG-2001; 2001US-0312147P.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
XX 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX 26-NOV-2001; 2001US-0333347P.
PA (GEHO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; AAA41726.
XX GENBANK; AAA41726.
PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX Example 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: the sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
```


SQ Sequence 547 AA;
Query Match 23.8%; Score 129.5; DB 7; Length 547;
Best Local Similarity 30.5%; Pred. No. 9.4e-06;
Matches 32; Conservative 27; Mismatches 41; Indels 5; Gaps 3;
QY 4 KSDEVFAKIAKRLSII-DPANRQVEHYVKFRITQG-GKVVQVWMDLKNVK---LVESDD 58
DB 432 KANLIFKEIEKKLEEGEFVKKIGGIFAFKVDGPGGKEATWVVDVKNKSGVLPDSDK 491
QY 59 AAEATLTWEDDIMEFAIGTALPAKEMAAQDKMEVDGQVELIFLLE 103
DB 492 KADCTITWADSDLLALMTGKNVPSQAFQGGKLTAGNGLAMKLQ 536
RESULT 4
ID ABB57301 standard; protein; 547 AA.
AC ABB57301;
DT 07-MAR-2002 (first entry)
DE Mouse ischaemic condition related protein sequence SEQ ID NO:843.
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
OS Mus musculus.
XN WO200188188-A2.
XX 22-NOV-2001.
XX 18-MAY-2001; 2001WO-JP004192.
PF 18-MAY-2000; 2000JP-00145977.
PR (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
PA Ishikawa K., Asai S, Takahashi Y, Nagata T, Ishii Y;
PI WPI; 2002-034733/04.
DR N-PSDB; ABI99762.
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX Claim 2; Page 2091-2093; 2690pp; English.
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (i) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (i). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX SQ Sequence 547 AA;
Query Match 23.6%; Score 128.5; DB 5; Length 547;
Best Local Similarity 31.4%; Pred. No. 1.2e-05;
Matches 33; Conservative 25; Mismatches 42; Indels 5; Gaps 3;

QY 4 KSDEVFAKIAKRLSII-DPANRQVEHYVKFRITQG-GKVVQVWMDLKNVK---LVESDD 58
DB 432 KANLIFKEIEKKLEEGEFVKKIGGIFAFKVDGPGGKEATWVVDVKNKSGVLPDSDK 491
QY 59 AAEATLTWEDDIMEFAIGTALPAKEMAAQDKMEVDGQVELIFLLE 103
DB 492 KADCTITWADSDLLALMTGKNVPSQAFQGGKLTAGNGLAMKLQ 536
RESULT 5
ID ADJ70149 standard; protein; 143 AA.
XX ADJ70149;
AC ADJ70149;
DT 06-MAY-2004 (first entry)
DE Human heat mitochondrial protein as a therapeutic target SeqID1955.
DE mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
OS Homo sapiens.
XX WO2003087768-A2.
XN 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010870.
PF 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX WPI; 2003-845369/78.
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX Claim 1; SEQ ID NO 1955; 180pp; English.
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nontropic, antidiabetic, and
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX SQ Sequence 143 AA;
Query Match 23.4%; Score 127.5; DB 7; Length 143;

Best Local Similarity	31.4%,	Pred. No. 2.8e-06;	
Matches	33;	Conservative	24; Mismatches 43; Indels 5; Gaps 3;
Qy	4	KSDEVFAKIAKRLSSI-DPANRQVHEVYVKPRITQG-KGVKKNWMDLKNVK---	LVSDD 58
Db	28	KANLVFEIEKKLBEEOQFKKIGGIFAFKVDGPGGKEATWVDVNGKGSVLPNSDK	87
Qy	59	AAEATLTMEDDIMPFAIGTALPAKEAMAQDRKMEVDGQVELIFLLE	103
Db	88	KADCTITMADSDFLALMTGKNPQSAFFQGLKITGNMGLAMKLIQ	132
RESULT 6			
ADJ71194			
ID	ADJ71194	standard; protein; 547 AA.	
XX	AC	ADJ71194;	
XX	DT	06-MAY-2004 (first entry)	
XX	DE	Human heat mitochondrial protein as a therapeutic target SeqID3000.	
XX	KW	mitochondrial; human; screening assay; diabetes mellitus;	
KW	KW	Huntington's disease; osteoarthritis;	
KW	KW	Leber's hereditary optic neuropathy; LHON;	
KW	KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;	
KW	KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;	
KW	KW	neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;	
KW	KW	osteopathic; ophthalmological; cytostatic.	
OS	OS	Homo sapiens.	
PN	PN	WO2003087768-A2.	
XX	PD	23-OCT-2003.	
XX	PF	04-APR-2003; 2003WO-US010870.	
XX	PR	12-APR-2002; 2002US-0372843P.	
PR	PR	17-JUN-2002; 2002US-039987P.	
XX	PR	20-SEP-2002; 2002US-0412418P.	
PA	PA	(MITO-) MITOKOR.	
PA	PA	(BUCK-) BUCK INST AGE RES.	
PI	PI	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;	
PI	PI	Warnock DE;	
XX	DR	WPI; 2003-845369/78.	
XX	PT	Identifying a mitochondrial target for drug screening assays and for	
PT	PT	treating diseases associated with altered mitochondrial function,	
PT	PT	comprises detecting a modified polypeptide in a sample and correlating	
XX	XX	with the disease.	
PS	PS	Claim 1; SEQ ID NO 3000; 180pp; English.	
XX	CC	This invention relates to novel mitochondrial targets that can be used	
CC	CC	for therapeutic intervention in treating a disease associated with	
CC	CC	altered mitochondrial function. Specifically, it refers to a method for	
CC	CC	identifying proteins of the human heart mitochondrial proteome that are	
CC	CC	useful for drug screening assays, as well as therapeutic targets. The	
CC	CC	present invention describes a method for identifying such proteins that	
CC	CC	can be used in the treatment of various diseases associated with altered	
CC	CC	mitochondrial function including diabetes mellitus, Huntington's disease,	
CC	CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial	
CC	CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy	
CC	CC	ragged red fibre syndrome (MERRF) or cancer. Accordingly, these	
CC	CC	compositions have neuroprotective, nontropic, antidiabetic,	
CC	CC	anticonvulsant, antiarthritic, osteopathic, ophthalmological and	
CC	CC	cytostatic activities. This polypeptide sequence is a human heart	
XX	XX	mitochondrial protein of the invention.	

CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 547 AA;

Query Match 23.4%; Score 127.5; DB 8; Length 547;
Best Local Similarity 31.4%; Pred. No. 1.6e-05;
Matches 33; Conservative 24; Mismatches 43; Indels 5; Gaps 3;

OY 4 KSDEVFAKIAKLESI-DPANQVEHYVYKFRITQG-GKVKVWVNDLKNV---LVESDD 58
DB 432 KANLVFKIEKLEEGEQVKIGGIFAFKVDGPGGKEATWVVDVKNRGKSVLPNSDK 491

OY 59 AAETLTMTMEDDIMPFAIGTGALPAKEMAAQDKMEVDGQVELIFLLE 103
DB 492 KADCTITWADSDFLALMTGKNPQSAFPQGLKITGNMGLAMKLQ 536

RESULT 8
AAW16329
ID AAW16329 standard; protein; 735 AA.

AC AAW16329;

DT 17-AUG-1997 (first entry)

DE Human host cell protein NSII-1.

XX NSII-1; non-structural protein 1 interactor 1; host cell protein;
KW influenza virus; replication; antiviral; virucide.

OS Homo sapiens.

FN WO9712967-A1.

PD 10-APR-1997.

PF 06-OCT-1995; 95WO-US013044.

PR 06-OCT-1995; 95WO-US013044.

XX (MOUN) MOUNT SINAI MEDICAL CENT.

PA Palest P, Oneill R;

PI WPI; 1997-226211/20.

XX N-PSDB; AAT63340.

PT New isolated DNA which encodes viral interacting proteins - used in
PT assays to isolate products for inhibiting viral protein binding which is
PT required for infection, replication, assembly or release.

XX Disclosure; Fig 12A-C; 98pp; English.

PS Non-structural protein 1 interactor 1 (NSII-1) (AAW16329) is a human host
CC cell protein which interacts with influenza virus protein NSI. It was
CC identified using a yeast interactive trap system. Its amino acid sequence
CC was deduced from NSII-1 cDNA (AAT63340). Another host cell protein, NPI-1
CC (AAW63327), has also been identified. These host cell proteins can be
CC used in assays to identify cpds. that interfere with the specific
CC interaction between the viral and host cell proteins. Such cpds. can be
CC used to treat viral infection

XX Sequence 735 AA;

Query Match 21.7%; Score 118.5; DB 2; Length 735;
Best Local Similarity 28.7%; Pred. No. 0.00025;
Matches 29; Conservative 25; Mismatches 42; Indels 5; Gaps 2;

OY 3 LKSDVEVFAKIAKRLSDIP-ANRQVEHYVYKFRITQGGKVKVWVNDLKN----VKLVESD 57
DB 432 KANLVFKIEKLEEGEQVKIGGIFAFKVDGPGGKEATWVVDVKNRGKSVLPNSDK 491

DB 621 LQSTFVEEIGRRRLKDIGPEVVKKNVAFWEHITKGNIGAKWTIDLKSGSGKYVQGPAP 680
OY 58 DAABATLTMTMEDDIMPFAIGTGALPAKEMAAQDKMEVDGQVEL 98
DB 681 GAUTTTILSDEDFMEVVLGKLDQPKAFFSGRLKARGNIML 721

RESULT 9

AAB70387

ID AAB70387 standard; protein; 735 AA.

AC AAB70387;

DT 02-MAY-2001 (first entry)

XX Human host cell protein NPII-1 protein sequence SEQ ID NO:13.

XX Identification; antiviral; viral protein; viral replication; NP;
KW viral infection; nucleoprotein.

OS Homo sapiens.

FN WO200111335-A2.

PD 15-FEB-2001.

PF 11-AUG-2000; 2000WO-US022257.

PR 11-AUG-1999; 99US-0148263P.

XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.

PI O'Neill R, Harty R, Palese PM;

DR WPI; 2001-168816/17.

DR N-PSDB; AAF59400.

PT Identifying a substance that inhibits the interaction between a viral
PT protein and a host cell protein, useful for the discovery of new
PT antiviral compounds.

PS Example; Fig 12; 147pp; English.

XX The present invention describes a method (M1) for identifying a substance
CC that inhibits the interaction of a viral protein (VP) with a host cell
CC protein (HP). The method comprises: (a) contacting HP with VP in the
CC presence of a test substance; and (b) detecting complex formation, where
CC the ability of the test substance to inhibit HP/VP interaction is
CC indicated by a decrease in complex formation. The antiviral compounds
CC that inhibit the interaction between a host protein (NSI-BP or NPI-1) and
CC a viral protein (NSI) are useful for treating or inhibiting viral
CC infection, preferably influenza and rhadovirus infection, in humans.
CC Antiviral compounds include peptides and antibodies. In particular
CC compositions comprising a polypeptide containing an amino acid sequence
CC corresponding to the NP-NLS domain of the influenza virus NP protein,
CC which inhibits the specific interaction of the NPI-1 protein with the
CC influenza virus NP protein are useful for treating or inhibiting
CC influenza viral infection in humans. The present sequence represents a
CC human host cell protein designated NPII-1, which is used in an example
CC from the present invention

XX Sequence 735 AA;

Query Match 21.7%; Score 118.5; DB 4; Length 735;
Best Local Similarity 28.7%; Pred. No. 0.00025;
Matches 29; Conservative 25; Mismatches 42; Indels 5; Gaps 2;

OY 3 LKSDVEVFAKIAKRLSDIP-ANRQVEHYVYKFRITQGGKVKVWVNDLKN----VKLVESD 57
DB 621 LQSTFVEEIGRRRLKDIGPEVVKKNVAFWEHITKGNIGAKWTIDLKSGSGKYVQGPAP 680

OY 58 DAABATLTMTMEDDIMPFAIGTGALPAKEMAAQDKMEVDGQVEL 98
DB 681 GAUTTTILSDEDFMEVVLGKLDQPKAFFSGRLKARGNIML 721

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 16:33:48 ; Search time 42 Seconds
(without alignments)
195.509 Million cell updates/sec

Title: US-10-823-203-3
Perfect score: 545
Sequence: 1 MSLKSDEVFAKIAKRLSIED.....EVDGQVELIFLEPFIAASK 110

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp.*
5: /cgn2_6/ptodata/1/iaa/PTCUS.COMB.pdp.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161.5	29.6	342	4	US-09-270-767-44008
2	127.5	23.4	289	4	US-09-949-016-6295
3	127.5	23.4	309	4	US-09-949-016-10025
4	127.5	23.4	547	4	US-09-538-092-984
5	118.5	21.7	735	4	US-09-636-791A-13
6	78.5	14.4	809	4	US-09-248-796A-19668
7	76	13.9	493	4	US-09-902-540-10842
8	75	13.8	462	4	US-09-252-991A-19002
9	69.5	12.8	404	4	US-09-902-540-12000
10	69	12.7	191	4	US-09-540-236-2322
11	69	12.7	294	3	US-08-137-175A-9
12	69	12.7	294	3	US-08-479-017-9
13	68.5	12.6	645	2	US-08-592-126-144
14	68.5	12.6	645	2	US-08-687-080-47
15	68.5	12.6	645	4	US-09-168-595-144
16	68.5	12.6	1312	2	US-08-592-126-148
17	68.5	12.6	1312	2	US-08-687-080-51
18	68.5	12.6	1312	4	US-09-168-595-148
19	67.5	12.4	322	4	US-09-583-110-3315
20	67.5	12.4	333	4	US-09-107-433-2981
21	67.5	12.4	821	4	US-09-543-681A-4722
22	67	12.3	290	4	US-09-248-796A-14651
23	66	12.1	307	4	US-09-270-767-45243
24	66	12.1	420	4	US-09-107-532A-5946
25	66	12.1	426	4	US-09-486-192-2
26	66	12.1	774	4	US-09-252-991A-29487
27	65.5	12.0	430	3	US-09-007-484-2

28	65.5	12.0	430	3	US-09-309-682-2	Sequence 2, Appli
29	65.5	12.0	2713	5	PCT-US96-01735-1	Sequence 1, Appli
30	65	11.9	399	4	US-09-270-767-48048	Sequence 48048, A
31	64.5	11.8	138	4	US-09-270-767-57049	Sequence 57049, A
32	64.5	11.8	498	1	US-08-470-202-59	Sequence 59, Appl
33	64.5	11.8	498	1	US-08-471-770-59	Sequence 59, Appl
34	64.5	11.8	498	2	US-08-468-059-59	Sequence 59, Appl
35	64.5	11.8	498	3	US-09-109-916-59	Sequence 59, Appl
36	64.5	11.8	498	4	US-09-886-156-59	Sequence 59, Appl
37	64.5	11.8	498	4	US-09-886-149-59	Sequence 59, Appl
38	64.5	11.8	498	4	US-09-886-150-59	Sequence 59, Appl
39	64.5	11.8	498	4	US-09-886-159-59	Sequence 59, Appl
40	64.5	11.8	498	4	US-10-328-090-59	Sequence 59, Appl
41	64.5	11.8	699	3	US-09-902-540-14095	Sequence 14095, A
42	64	11.7	393	4	US-09-308-003-18	Sequence 18, Appl
43	64	11.7	515	4	US-09-248-796A-19787	Sequence 19787, A
44	64	11.7	1225	4	US-09-949-016-9468	Sequence 9468, Ap
45	63.5	11.7	273	4	US-09-248-796A-15037	Sequence 15037, A

ALIGNMENTS

RESULT 1

US-09-270-767-44008
; Sequence 44008, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44008.
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44008

Query Match	29.6%	Score	161.5	DB	4	Length	342
Best Local Similarity	34.9%	Pred. No.	5e-12				
Matches	38	Conservative	19	Mismatches	49	Indels	3
Gaps	1						
Qy	1	MSLKSDEVFAKIAKRLSIEDPANQVEHYKFRITOGKVVNWMDLKNVLFVSD--	58				
Db	48	MSLQSDAVFQKIIDGLKENEAKAVNGVFLYKITKDGKAKVETLDCNKAKEYGPAQG	107				
Qy	59	-AAEATLTMEDDIMPATGTGALPAKEMAQDKMEVDGQVELIFLEPFI	106				
Db	108	IKVDTTLTVADEDMWDIALGKLNFOAPFMKGLKIAGNIMLTQKLAPLL	156				
RESULT 2							
US-09-949-016-6295							
; Sequence 6295, Application US/09949016							
; Patent No. 6812339							
; GENERAL INFORMATION:							
; APPLICANT: VENTER, J. Craig et al.							
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED							
; FILE REFERENCE: CL001307							
; CURRENT APPLICATION NUMBER: US/09/949,016							
; CURRENT FILING DATE: 2000-04-14							
; PRIOR APPLICATION NUMBER: 60/241,755							
; PRIOR FILING DATE: 2000-10-20							
; PRIOR APPLICATION NUMBER: 60/237,768							
; PRIOR FILING DATE: 2000-10-03							
; PRIOR APPLICATION NUMBER: 60/231,498							
; PRIOR FILING DATE: 2000-09-08							


```

RESULT 10
US-09-540-236-2322
; Sequence 2322, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2322
; LENGTH: 191
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2322

```

```

1  APPLICANT: Gary L. Breton et al.
2  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
3  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
4  FILE REFERENCE: 2709.2005-001
5  CURRENT APPLICATION NUMBER: US/09/540.236
6  CURRENT FILING DATE: 2000-04-04
7  NUMBER OF SEQ ID NOS: 3840
8  SEQ ID NO 2322
9  LENGTH: 131
10 TYPE: PRT
11 ORGANISM: M.catarrhalis
12 US-09-540-236-2322

```


RESULT 14
US-08-687-080-47
; Sequence 47, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:

```

RESULT 15
US-09-168-595-144
; Sequence 144, Application US/09168595
; Patent No. 655566
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/168,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,126
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 144:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 645 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: G18.pap
US-09-168-595-144

Query Match      12.6%; Score 68.5; DB 4; Length 645;
Best Local Similarity 26.8%; Pred. No. 7.7;
Matches 30; Conservative 23; Mismatches 32; Indels 27; Gaps 6;

QY      4 KSDEVFA--KIAKRLSIPANRQVEHVYKFRITQGGKVVKNWVMDLKNVK-----L 53
Db      175 KFDEIFSATRYIKALETL-----RQV-----RQTQGQK-VKEYQMLKYLKQYKEKACEI 223
QY      54 VESDDAAEATLTWEDDIMPFAIGTGALPAKEAQAQDK-----MEVDGQVELI 99
Db      224 RQITSKEAQLTSSKEIVKSYENELDPLKNRLKEIEHNL SKIMKLDNEIKAL 275

Search completed: February 28, 2005, 16:41:36
Job time : 43 secs
```

Fried. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.


```
Query Match      15.2%; Score 83; DB 16; Length 345;
Best Local Similarity 26.0%; Pred. No. 0.53;
Matches 27; Conservative 17; Mismatches 48; Indels 12; Gaps 4;

Qy      6 DEVFAKIAKRLESIDPANRQVHVYKFRIT--QGKVVKNVMDLK-----NVKLVESDDA 59
Db      238 EETFRIVKDSLS--DDVVKATQAIYLPFELSGEDGG---TWFLDLKSKGNGVGYGEPDQ 291
Qy      60 AEATLTMEDDIMPATGTCALPAKEAQAQDKMEVDGQVELIFLLE 103
Db      292 ADVVMSMTTDDFVRKMFSGKLKPTMAFMGSKLKIKGNMALAIKLE 335

RESULT 13
US-09-796-089-7
; Sequence 7, Application US/09796089
; Patent No. US20020010946A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 21612, 21615, 21620, 21676, 33756, No. US20020010946A1e1
; FILE REFERENCE: 35800/208222
; CURRENT APPLICATION NUMBER: US/09/796,089
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/33873
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/464,039
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-089-7

Query Match      15.2%; Score 83; DB 9; Length 418;
Best Local Similarity 26.0%; Pred. No. 0.69;
Matches 27; Conservative 17; Mismatches 48; Indels 12; Gaps 4;

Qy      6 DEVFAKIAKRLESIDPANRQVHVYKFRIT--QGKVVKNVMDLK-----NVKLVESDDA 59
Db      311 EETFRIVKDSLS--DDVVKATQAIYLPFELSGEDGG---TWFLDLKSKGNGVGYGEPDQ 364
Qy      60 AEATLTMEDDIMPATGTCALPAKEAQAQDKMEVDGQVELIFLLE 103
Db      365 ADVVMSMTTDDFVRKMFSGKLKPTMAFMGSKLKIKGNMALAIKLE 408

RESULT 14
US-10-092-900A-242
; Sequence 242, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spyttek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
```

```
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieser, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; REMAINING Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 242
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-242

Query Match      15.2%; Score 83; DB 15; Length 418;
Best Local Similarity 26.0%; Pred. No. 0.69;
Matches 27; Conservative 17; Mismatches 48; Indels 12; Gaps 4;

Qy      6 DEVFAKIAKRLESIDPANRQVHVYKFRIT--QGKVVKNVMDLK-----NVKLVESDDA 59
Db      311 EETFRIVKDSLS--DDVVKATQAIYLPFELSGEDGG---TWFLDLKSKGNGVGYGEPDQ 364
Qy      60 AEATLTMEDDIMPATGTCALPAKEAQAQDKMEVDGQVELIFLLE 103
Db      365 ADVVMSMTTDDFVRKMFSGKLKPTMAFMGSKLKIKGNMALAIKLE 408

RESULT 15
US-10-375-039-34
; Sequence 34, Application US/10375039
; Publication No. US20040170986A1
; GENERAL INFORMATION:
; APPLICANT: USUDA, Yoshihiro
; APPLICANT: NISHIO, Yosuke
; APPLICANT: YASUEDA, Hisashi
; APPLICANT: SUGIMOTO, Shinichi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING POLYPEPTIDES INVOLVED IN AMINO ACID BIOS
; FILE REFERENCE: 232743USO
; CURRENT APPLICATION NUMBER: US/10/375,039
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 54
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 16:25:38 ; Search time 39 Seconds
(without alignments)
271.381 Million cell updates/sec

Title: US-10-823-203-3
Perfect score: 545
Sequence: 1 MSLKSDVEFAKIAKRLSID.....EVDGQVELIFLLEPFIA SLK 110

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129.5	23.8	547	2 S34744	sterol carrier pro
2	129.5	23.8	547	2 A39368	sterol carrier pro
3	128.5	23.6	547	2 JU0157	sterol carrier pro
4	127.5	23.4	547	2 I38205	sterol carrier pro
5	126.5	23.2	143	2 A40015	sterol carrier pro
6	118.5	21.7	736	2 S59136	estradiol 17beta-d
7	113.5	20.8	735	2 S74209	multifunctional be
8	109.5	20.1	546	2 B40407	sterol carrier pro
9	104.5	19.2	118	2 T28068	hypothetical prote
10	100	18.3	436	2 T16638	hypothetical prote
11	75	13.8	454	2 A82952	glucosamine-1-phos
12	74	13.6	241	2 B71237	hypothetical prote
13	74	13.6	254	1 FRPBH	ferritin heavy cha
14	73.5	13.5	291	2 T10966	6-phosphogluconate
15	73	13.4	190	2 E89803	conserved hypothet
16	73	13.4	427	2 E87669	conserved hypothet
17	73	13.4	528	2 C69323	conserved hypothet
18	72	13.2	298	2 S50964	hypothetical prote
19	72	13.2	492	2 AC2781	trigger factor (im
20	72	13.2	497	2 D97560	trigger factor (tf
21	72	13.2	653	2 H82316	conserved hypothet
22	71	13.0	250	2 A40992	ferritin precursor
23	71	13.0	393	2 AC1142	N-acyl-L-amino aci
24	71	13.0	430	2 H81389	thiamin biosynthes
25	70.5	12.9	1579	2 S59801	protein kinase SSK
26	70	12.8	293	2 T35157	6-phosphogluconate
27	70	12.8	1333	2 E84601	probable retroelem
28	69.5	12.8	458	2 AH0029	trk system potassi
29	69.5	12.8	904	2 C70559	probable polA prot

30 69 12.7 180 2 F97149 probable flavodoxi
31 69 12.7 1642 2 T19130 hypothetical prote
32 68.5 12.6 323 2 C83940 sugar ABC transpor
33 68 12.5 134 2 S17647 NADH2 dehydrogenas
34 67.5 12.4 137 2 C81737 conserved hypothet
35 67.5 12.4 321 2 A95182 hypothetical prote
36 67.5 12.4 322 2 D98049 thiorodoxin-disulf
37 67.5 12.4 1302 1 JC6009 surface-located me
38 67 12.3 234 2 F71845 mucoate cycloisom
39 67 12.3 345 2 H72429 hypothetical prote
40 67 12.3 673 2 T05619 probable homeodoma
41 67 12.3 689 2 A85295 hypothetical prote
42 67 12.3 781 1 Q0BEN7 helicase (EC 3.6.1
43 67 12.3 1232 2 D64413 cobalamin biosynth
44 66.5 12.2 805 2 E70474 translation initia
45 66.5 12.2 1529 2 T20986 hypothetical prote

ALIGNMENTS

RESULT 1
S34744
sterol carrier protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S34744
R:Pfeifer, S.M.; Sakuragi, N.; Ryan, A.; Johnson, A.L.; Deeley, R.G.; Billheimer, J.T.; E
Arch. Biochem. Biophys. 304, 287-293, 1993
A>Title: Chicken sterol carrier protein 2/sterol carrier protein x: cDNA cloning reveals
A:Reference number: S34744; MUID:93312016; PMID:8323294
A:Accession: S34744
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-547 <PE>
A:Cross-references: UNIPROT:Q07598; GB:L09231; NID:g304422; PIDN:AAA02488.1; PID:g304423

Query Match 23.8%; Score 129.5; DB 2; Length 547;
Best Local Similarity 33.3%; Pred. No. 5.9e-05;
Matches 35; Conservative 22; Mismatches 43; Indels 5; Gaps 3;

QY 4 KSDEVFAKIAKRL-ESIDPANQVEHYVKFRTQG-GKVKVNWMDLKNVK---LVESDD 58
DB 432 KSHLVFKEIEKQLQEGEQFVKIKGVPFAFKIDGPGGKEATWVVDVKNKGSAVNSDK 491
QY 59 AAEATLTMEDDIMEAIGTGALPAKEAMADKMEVDGQVELIFLLE 103
DB 492 KADCTITWADTDLALMTGKNVPTAFQGGKLKISGNMGMAKQLQ 536

RESULT 2
A39368
sterol carrier protein 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: A39368; A39054; B39054; A29366; A27661; S17842; A34584; A34635
R:Mori, T.; Tsukamoto, T.; Mori, H.; Tashiro, Y.; Fujiki, Y.
Proc. Natl. Acad. Sci. U.S.A. 88, 4338-4342, 1991
A>Title: Molecular cloning and deduced amino acid sequence of non-specific lipid transfer
sequence of nonspecific lipid transfer protein as its C-terminal part.
A:Reference number: A39368; MUID:91239563; PMID:2034675
A:Accession: A39368
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-547 <MOR>
A:Cross-references: UNIPROT:P11915; GB:M62763; NID:g202552; PIDN:AAA40622.1; PID:g202553
J. Biol. Chem. 266, 630-636, 1991
A>Title: Cloning, expression, and nucleotide sequence of rat liver sterol carrier protein
A:Reference number: A39054; MUID:91093192; PMID:1985920
A:Accession: A39054
A>Status: preliminary
A:Molecule type: mRNA

Db 73 KVLKSVLDSVKGLNLNVEVYVVRVDGNDVFWLVSFLRERVRVNGALVT---EGV 129
Qy 70 IMFALGTGALPAKAMAQDKMEVDGQVELIFL-----LEPIASLK 110
Db 130 KIYPGGMGNLTNEFPVREKEL--ELSLAYLAKLDGILEKYRGSWR 174

RESULT 13

FRFBH

ferritin heavy chain precursor - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S17426; S14868
R:Spence, M.J.; Henzl, M.T.; Lammers, P.J.
Plant Mol. Biol. 17, 499-504, 1991
A:Title: The structure of a Phaseolus vulgaris cDNA encoding the iron storage protein ferritin
A:Reference number: S17426; MUID:91355941; PMID:1884000
A:Accession: S17426
A:Molecule type: mRNA
A:Residues: 1-254 <SP>
C:Cross-references: UNIPROT:P25699; EMBL:X58274; NID:G21026; PIDN:CAA41213.1; PID:G21027
R:Spence, M.J.; Sengupta-Gopalan, C.; Henzl, M.T.; Lammers, P.J.
submitted to the EMBL Data Library, March 1991
A:Description: A Phaseolus vulgaris cDNA encoding an iron storage protein with sequence
A:Reference number: S14868
A:Accession: S14868
A:Molecule type: mRNA
A:Residues: 1-69, 'X', 71-125, 'X', 127-254 <SP2>
C:Cross-references: EMBL:X58274; NID:G21026; PIDN:CAA41213.1; PID:G21027
C:Genetics:
A:Gene: pfe
C:Superfamily: ferritin
C:Keywords: chloroplast; iron; iron storage; metalloprotein
P:1-48/Domain: transit peptide (chloroplast) #status predicted <TNP>
P:49-254/Product: ferritin heavy chain #status predicted <WAT>
F:199,133,134,136,137,183/Binding site: iron (Glu, Glu, Glu, His, Glu) #status predicted

Query Match 13.6%; Score 74; DB 1; Length 254;
Best Local Similarity 25.9%; Pred. No. 7;
Matches 29; Conservative 28; Mismatches 43; Indels 12; Gaps 4;
Qy 9 FAKIAKRLSSIDPANRQVEHVYKFRITOGGKVV-----KNWMDLKNVKLVESDDAAEATL 64
Db 123 FARFPK--ESSEEREHAELKMYQTRGGRVLPKIKNPSEPEHVEKGDALYAMELAL 180
Qy 65 TME---DDIMFAIGTGALPAKAMAQDKME---VDGQVELIFLLEPFTIASLK 110
Db 181 SLEKLVNEKLSRVSHVADRNKDPQLADFTIESEFLSEQVEAIKKISEYVAQR 232

RESULT 14

T10966

6-phosphogluconate dehydrogenase homolog - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10966
R:Calcutt, M.J.
Gene 151, 23-28, 1994
A:Title: Gene organization in the dnaA-gyrA region of the Streptomyces coelicolor chromosome
A:Reference number: Z17235; MUID:95129865; PMID:7828880
A:Accession: T10966
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-291 <CAL>
C:Cross-references: UNIPROT:Q53917; EMBL:L27063; NID:g436023; PID:g436024
A:Experimental source: strain A3(2)
C:Superfamily: Aquifex aeolicus 6-phosphogluconate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase
F:3-360/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>

Query Match 13.5%; Score 73.5; DB 2; Length 291;
Best Local Similarity 29.6%; Pred. No. 9.2;
Matches 29; Conservative 19; Mismatches 29; Indels 21; Gaps 5;

Qy 7 EVFAKIAKRLSSIDPANRQVEHVYK-FRITOGGKVVKNWMDLKNVKLVESD----- 57
Db 176 QAYAEGWELLEKVD-----SVENVREVFRSWQEGTVIRSWLLDLA-VNALDDDEHLDGLRG 230
Qy 58 ---DAAEATLTMEDDIMEAIG-----TGALPAKAMAQD 88
Db 231 YAQDSGEGRWTVAAIDNAVPLPAITASLFAFASRQD 268

RESULT 15

E89803

conserved hypothetical protein SA0359 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89803
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Iji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <KUR>
C:Cross-references: UNIPROT:Q99WK5; GB:BA000018; PID:g13700286; PIDN:BAB41584.1; GSPDB:G1
C:Experimental source: strain N315
C:Genetics:
A:Gene: SA0359

Query Match 13.4%; Score 73; DB 2; Length 190;
Best Local Similarity 24.5%; Pred. No. 6.3;
Matches 23; Conservative 22; Mismatches 43; Indels 6; Gaps 3;
Qy 4 KSDEVFAKIAKRLSSIDPANRQVEHVYKFRITOGGKVVKNWMDLKNVKLVESDDAAEAT 63
Db 61 KAETTYK--GOKLKGISFENSGEMWAYKVTOOKGSESEVLVAD-KNKKVINKKTEKPT 117
Qy 64 LTMEDDIMEAIGTGALPAKAMAQDKMEVDGQVE 97
Db 118 VNENDNPKY---SDAIDYKKAISKQKEFDGDIK 148

Search completed: February 28, 2005, 16:40:49

Job time : 41 secs

IMAGE IS BLANK